



全国统计教材编审委员会“十二五”规划教材

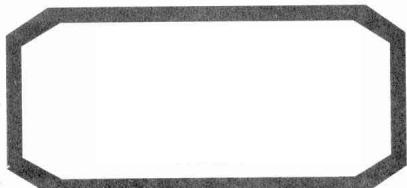
抽样调查理论与方法

第二版

冯士雍 倪加勋 邹国华 编著



中国统计出版社
China Statistics Press



全国统计教材
“十一五”规划教材

抽样调查理论与方法

第二版

冯士雍 倪加勋 邹国华 编著

(京)新登字 041 号

图书在版编目(CIP)数据

抽样调查理论与方法 / 冯士雍, 倪加勋, 邹国华编著. —2 版. — 北京:
中国统计出版社, 2012. 8

全国统计教材编审委员会“十二五”规划教材

ISBN 978-7-5037-6592-6

I. ①抽… II. ①冯… ②倪… ③邹… III. ①抽样调查统计—
高等学校—教材 IV. ①C811

中国版本图书馆 CIP 数据核字(2012)第 167116 号

抽样调查理论与方法(第二版)

作 者/冯士雍 倪加勋 邹国华

责任编辑/杨映霜 梁 超

封面设计/上智博文

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 57 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

网 址/<http://csp.stats.gov.cn>

电 话/邮购(010)63376907 书店(010)68783172

印 刷/三河双峰印刷装订有限公司

经 销/新华书店

开 本/710×1000mm 1/16

字 数/380 千字

印 张/24.5

印 数/1—3000 册

版 别/2012 年 9 月第 1 版

版 次/2012 年 9 月第 1 次印刷

书 号/ISBN 978-7-5037-6592-6/C.2676

定 价/48.00 元

版权所有。未经许可,本书的任何部分不得以任何方式在世界任何地区
以任何文字翻印、拷贝、仿制或转载。

中国统计版图书,如有印装错误,本社发行部负责调换。

全国统计教材编审委员会

顾问 罗 兰 袁 卫 冯士雍 吴喜之
方积乾 王吉利 庞 皓 李子奈

主任 徐一帆

副主任 严建辉 田鲁生 邱 东 施建军
耿 直 徐勇勇

委员(按姓氏笔划排序)

丁立宏	万崇华	马 骏	毛有丰	王兆军
王佐仁	王振龙	王惠文	丘京南	史代敏
龙 玲	刘建平	刘俊昌	向书坚	孙秋碧
朱 胜	朱仲义	许 鹏	余华银	张小斐
张仲梁	张忠占	李 康	李兴绪	李宝瑜
李金昌	李朝鲜	杨 虎	杨汭华	杨映霜
汪荣明	肖红叶	苏为华	陈 峰	陈相成
房祥忠	林金官	罗良清	郑 明	柯惠新
柳 青	胡太忠	贺 佳	赵彦云	赵耐青
凌 兮	唐年胜	徐天和	徐国祥	郭建华
崔恒建	傅德印	景学安	曾五一	程维虎
蒋 萍	潘 瑶	颜 虹		

出版说明

“十二五”时期，是我国全面实施素质教育，全面提高高等教育质量，深化教育体制改革，推动教育事业科学发展，提高教育现代化水平的时期。“十二五”伊始，统计学迎来了历史性的重大变革和飞跃。2011年2月，在国务院学位委员会第28次会议通过的新的《学位授予和人才培养学科目录(2011)》(以下简称“学科目录”)中，统计学从数学和经济学中独立出来，成为一级学科。这一变革和飞跃将对中国统计教育事业产生巨大而深远的影响，中国统计教育事业将在“十二五”时期发生积极变化。

正是在这一背景下，全国统计教材编审委员会制定了《“十二五”全国统计教材建设规划》(以下简称“规划”)。根据“学科目录”在统计学下设有数理统计学，社会经济统计学，生物卫生统计学，金融统计、风险管理与精算学，应用统计5个二级学科的构架，“规划”对“十二五”全国统计规划教材建设作了全面部署，具有以下特点：

第一，打破以往统计规划教材出版学科单一的格局。全面发展数理统计学，社会经济统计学，生物卫生统计学，金融统计、风险管理与精算学，应用统计5个二级学科规划教材的出版，使“十二五”全国统计规划教材涵盖5个二级学科，形成学科全面并平衡发展的出版局面。

第二，打破以往统计规划教材出版层次单一的格局。在编写出版好各学科本科生教材的基础上，对研究生教材出版进行深入研究，出版一批高水平高层次的研究生教材，为我国研究生教育、尤其是应用统计研究生教育提供教学服务。同时，积极重视统计专科教材出版，联合各专科院校，组织编写和出版适应统计专科教学和学习的优秀教材。

第三，打破以往统计规划教材出版品种单一的格局。鼓励内容创新，联系统计实践，具有教学内容和教学方法特色的、各高校自编的相同内容选题的精品教材出版，促进统计教学向创新性、创造性和多样性

发展。

第四,重视非统计专业的统计教材出版。探讨对非统计专业学生的统计教学问题,为非统计专业学生组织编写和出版概念准确、叙述简练、深入浅出、表达方式活泼、练习题贴近社会生活的统计教材,使统计思想和统计理念深入非统计专业学生,以达到统计教学的最大效果。

第五,重视配合教师教学使用的电子课件和辅助学生学习使用的电子产品的配套出版,促进高校统计教学电子化建设,以期最后能形成系统,提高统计教育现代化水平。

第六,重视对已经出版的统计规划教材的培育和提高,本着去粗存精、去旧加新、与时俱进的原则,继续优化已经出版的统计教材的内容和写作,强化配套课件和习题解答,使它们成为精品,最后锤炼成为经典。

“十二五”期间,编审委员会将本着“重质量,求创新,出精品,育经典”的宗旨,组织我国统计教育界专家学者,编写和编辑出版好本轮教材。本轮教材出版后,将能够形成学科齐全、层次分明、品种多样、配套系统的高质量立体式结构,使我国统计规划教材建设再上新台阶,这将对推动我国统计教育和统计教材改革,推动我国统计教育事业发展,提高我国统计教育现代化水平产生积极意义。

让教师的教学和学生的学习事半功倍,并使学生在毕业之后能够学以致用的统计教材,是本轮教材的追求。编审委员会将努力使本轮教材好教、好学、好用,尽力使它们在内容上和形式上都向国外先进统计教材看齐。限于水平和经验,在教材的编写和编辑出版过程中仍会有不足,恳请广大师生和社会读者提出批评和建议,我们将虚心接受,并诚挚感谢!

全国统计教材编审委员会

2012年7月

第二版前言

近年来,抽样调查已成为所有统计学专业的一门核心课程。不同背景的统计学专业的本科生、硕士研究生、博士研究生都将抽样调查作为一门必修课,也有越来越多的管理科学、社会学、传媒学以及许多从事自然科学及人文科学的,特别是一些实验科学专业的学生也选修了这门课程。

本书第一版自 1998 年出版以来受到了各方面的关注和欢迎,共进行了四次印刷,许多院校将本书作为教材使用。考虑到十几年来抽样调查的理论与应用的发展以及不同学校对这门课程教学安排的实际情况,有必要对本书进行修订,以更好地满足各方面的需要。

本书的第二版对第一版作了全面的修订及增补,主要变化有:

1)前十二章的内容沿用了原有的体例。除改正文字、计算及印刷错误外,增加了对一些结果的统计思想的解释,改进或简化了某些公式的推导。

2)在第一章概论中补充了抽样调查理论及应用的历史特别是近年来的发展;原第六章二重抽样中增加了为 PPS 抽样的二重抽样等内容,第七章不等概率抽样中增加了泊松抽样,同时对调了两章的顺序;在第十一章中增加了数量化特征敏感性问题的随机化回答技术等内容。

3)新增加的第十三章“抽样调查的推断框架及小域估计”介绍了抽样调查的三种基本推断框架,这既总结了经典抽样调查推断即基于设计的推断,也引进了基于模型的推断及模型辅助推断,后两种推断是现代抽样调查理论及方法研究的重要手段;同时也介绍了在应用中非常重要的小域估计方法。这一章对那些想了解这门学科研究与应用前沿的学生是十分有用的。

4)新增加的第十四章“案例分析”是根据编著者参与的三个全国性的实际抽样调查项目的技术文件改写的。每个案例都交代了项目的背景,针对项目具体要求,着重讨论了所采用的抽样设计(包括相应的数据处理方法),并进行了简短的讨论。目的是使读者对实际项目的抽样设计的各个

重要环节有较为系统的了解。

5)增加了“部分习题答案与提示”，供学生和任教老师参考。

为使本教程能尽可能满足各方面的需要，我们既强调概念及方法的统计意义和思想，又兼顾科学的严谨性及实际操作的可行性。在介绍具体抽样方法时，对每种方法的适用条件及实施程序都作了详尽的说明，对不同方法进行了比较；对重要的结果给出了严格的数学证明；同时精心选编了数值例子，其中相当部分选自我们实际接触到的案例。不过需要指出的是，正因为这些例子大多是有实际背景的，随着时间的推延，具体数值有些已与当前情况有一定的差距，但相信这不会给学生的理解带来实质影响。

由于本教材包含的内容较多，难度相差也较大，在实际使用时应根据教学对象及教学计划作适当安排。总的说来，前十一章的内容（打“*”号的节除外）适合大部分本科生的教学，不过对那些数学要求不很强的院校，可以略去某些较困难的数学推导；第十二章、第十三章适合于研究生教学；第十四章则更可以根据不同读者需要选用。

本书的修订最早在 2007 年就启动，倪加勋教授提供了 7.3 节“为 PPS 抽样的二重抽样”的初稿及部分习题的参考答案；第二版的其余修订工作主要是我们两人完成的，所以我们对本书中可能存在的问题与错误负责。

在修订过程中，得到了曾采用本书为教材的许多院校老师的帮助，他们指出了第一版存在的错误，提出了许多宝贵的修改建议。借此机会，向他们表示诚挚的谢意。我们特别感谢如下为本书的修订付出辛勤劳动的老师和学生：李莉莉、丁文兴、涂玉娟、杨军、李锋、邓国华等提出了相当具体的建议；李锋还将第一版的内容输入成电子稿；王海鹰、朱荣对习题答案作了全面的核对并给出了部分证明的提示；陈心洁审读了第二版的大部分初稿。另外在修订过程中得到了中国统计出版社严建辉社长、陈悟朝主任等同志的热情鼓励，本书的责任编辑杨映霜、梁超等细致、出色的工作也为保证本书的质量作出了贡献，特此表示衷心的感谢！

冯士雍 邹国华

于中国科学院数学与系统科学研究院

2012 年 7 月

第一版前言

抽样调查是统计学的一门重要分支学科。随着我国经济体制从计划经济体制向社会主义市场经济体制的转变,抽样调查的作用显得日益重要。我国统计调查方法改革目标模式已明确我国统计调查“以经常性抽样调查为主体”,并列入修改后的《统计法》。在国际上,抽样调查早已成为世界各国普遍采用的一种统计调查方法。它的理论和方法研究发展很快。但是在我国统计学专业中,长期以来还缺少一本适合社会发展需要,建立在严格科学基础上的抽样调查教材。本书是根据1993年6月全国教材编审委员会讨论通过的《抽样调查理论与方法》教学大纲编写,作为高等院校统计学专业规划教材之一出版的。

根据大纲要求,本课程的教学目的是使学生能系统正确地掌握抽样调查的理论基础,掌握几种基本的抽样方法,培养学生初步具有能结合实际情况对具体项目进行抽样设计和对所获得数据进行处理和分析的能力。为此,在教学中要特别讲清各种概念和方法的统计思想,使学生牢固地掌握基本抽样技术的实施和分析方法,并通过基本训练有进一步学习文献、扩大知识面的能力。

本书共分十二章。第一章为概论,简述抽样调查的意义与作用,它的发展简史与主要应用。第二章是有关抽样的基本概念,为以后各章的学习打好基础。第三章至第十章是本课程的主体,逐章介绍简单随机抽样、分层随机抽样、比估计与回归估计、二重抽样、不等概率抽样、整群抽样、二阶与多阶抽样及系统抽样。详细介绍这些具体抽样方法的适用场合与实施方法、对总体目标量的估计、估计量的方差与方差估计。第十一章介绍非抽样误差。这些都是大纲规定的内容(某些超出大纲要求的节或小节以“*”号表示)。原大纲的第十二章是“案例分析”,但考虑到案例分析要占

大量篇幅，在实际教学中也应根据教师与学生不同的需要而定，不宜作硬性统一规定，因此全国统计教材编审委员会考虑这方面内容以后另行编写出版。由于复杂抽样方差估计在现代抽样理论与实际应用的重要性，我们在最后一章中增加了这方面的内容。但这并非大纲所要求的，只作为补充教材使用。

本书的第二、三、四、五、七、八、九诸章由中国科学院系统科学所冯士雍研究员执笔并任主编；第一、六、十、十一诸章由中国人民大学统计学系倪加勋教授提出初稿并经冯士雍研究员增删与改写；第十二章则由中国科学院系统科学所邹国华博士执笔。全书习题主要由倪加勋教授提供并经邹国华博士补充。书中的大多数例子都是经编著者精选的有实际背景的数值例子，个别例子（包括部分习题）选自所列参考书目，未一一注明，特向有关作者和出版社表示感谢。需要说明的是，由于篇幅所限有相当部分的数值例子的样本量都较小，这对实际应用是不够的。这些例子只能作为说明性的示例，请读者务必不要照搬。

北京大学概率统计系孙山泽教授对全书的手稿进行了仔细的审阅并提出许多宝贵意见，为本书增色不少。国家统计局统计教育中心与中国统计出版社的有关同志对本书的编写与出版给予了大力支持、鼓励与帮助，在此一并表示深深的感谢。

本书在编写过程中尽管对内容与结构经过反复的斟酌与多次修改，曾几易其稿，但由于我们学术水平与教学经验所限，书中错误及选材结构安排不当之处在所难免，恳请各方面专家及使用本教材的老师与同学们不吝批评指正，以便在再版时修改与完善。

编著者

1997年8月

目 录

第一章 概论	(1)
1.1 抽样调查的意义与作用	(1)
1.1.1 什么是抽样调查	(1)
1.1.2 非概率抽样	(2)
1.1.3 概率抽样	(3)
1.1.4 抽样调查的作用	(4)
1.1.5 抽样调查与普查的关系	(5)
1.2 抽样调查的发展历史	(6)
1.2.1 国际抽样调查的发展	(6)
1.2.2 我国抽样调查的发展	(10)
1.3 抽样调查的主要应用	(12)
习题	(14)
 第二章 基本概念	(16)
2.1 总体与样本	(16)
2.1.1 目标总体与抽样总体	(16)
2.1.2 抽样单元与抽样框	(17)
2.1.3 总体指标的类型	(18)
2.1.4 样本、样本量与抽样比	(19)
* 2.1.5 固定总体模型与超总体模型	(20)
2.2 几种基本的抽样方法	(20)
2.2.1 简单随机抽样	(20)
2.2.2 分层抽样	(21)
2.2.3 整群抽样	(21)
2.2.4 多阶抽样	(21)
2.2.5 系统抽样	(22)

2.2.6 不等概率抽样	(22)
2.3 误差与精度的表示方法	(23)
2.3.1 抽样调查中的误差来源	(23)
2.3.2 均方误差、方差与偏倚	(23)
2.3.3 误差限与置信度	(25)
2.3.4 精度与费用	(27)
习题	(27)

第三章 简单随机抽样 (29)

3.1 概述	(29)
3.1.1 什么是简单随机抽样	(29)
3.1.2 简单随机抽样的实施方法	(30)
3.1.3 简单随机抽样在抽样理论中的地位与作用	(34)
3.2 总体均值与总量的简单估计	(34)
3.2.1 简单估计量及其无偏性	(34)
3.2.2 估计量的方差	(37)
3.2.3 估计量的方差估计	(40)
3.3 总体比例的简单估计	(42)
3.3.1 对总体的描述	(42)
3.3.2 估计量及其性质	(43)
3.4 样本量的确定	(45)
3.4.1 确定样本量的原则与主要考虑因素	(45)
3.4.2 估计总体均值或总量时样本量的确定方法	(46)
3.4.3 估计总体比例时样本量的确定方法	(48)
* 3.4.4 逆抽样方法	(49)
3.5 放回简单随机抽样	(51)
3.5.1 估计量及其性质	(51)
3.5.2 设计效应与样本量的确定	(53)
* 3.6 子总体的估计	(54)
3.6.1 问题的提出	(54)
3.6.2 子总体均值的估计	(55)
3.6.3 子总体总量的估计	(55)

习题	(58)
第四章 分层随机抽样	(61)
4.1 概述	(61)
4.1.1 什么是分层抽样和分层随机抽样	(61)
4.1.2 分层抽样的特点和适用场合	(61)
4.1.3 记号	(62)
4.2 简单估计量及其性质	(63)
4.2.1 对总体均值或总量的估计	(63)
4.2.2 对总体比例的估计	(65)
4.3 各层样本量的分配	(67)
4.3.1 样本量分配对精度的影响	(67)
4.3.2 比例分配	(68)
4.3.3 最优分配	(69)
* 4.3.4 某些层需要超过 100% 抽样时的修正	(72)
4.4 样本总量的确定	(74)
4.4.1 影响样本总量的因素	(74)
4.4.2 估计总体均值的情形	(75)
4.4.3 估计总体总量的情形	(77)
4.4.4 估计总体比例的情形	(78)
4.4.5 给定总费用时样本量的确定方法	(79)
4.5 分层随机抽样效果分析	(80)
4.5.1 分层随机抽样与简单随机抽样的比较	(80)
4.5.2 最优分配在精度上的改进	(82)
4.5.3 估计总体比例时分层和最优分配精度上的得益	(83)
4.5.4 从分层样本估计分层抽样的效果	(83)
* 4.5.5 分层标志的选择及最优分层	(86)
* 4.5.6 层数的确定	(88)
* 4.6 若干进一步问题	(89)
4.6.1 偏离最优分配时对方差的影响	(89)
4.6.2 多指标情形样本量的分配	(90)
4.6.3 层权误差对估计量的影响	(92)

4.6.4 事后分层	(93)
习题	(96)
第五章 比估计与回归估计	(101)
5.1 概述	(101)
5.1.1 问题的提出	(101)
5.1.2 比估计与回归估计的作用与使用条件	(102)
5.2 比估计	(102)
5.2.1 定义及基本性质	(102)
5.2.2 方差估计及置信限	(107)
5.2.3 比估计与简单估计的比较	(112)
* 5.2.4 乘积估计	(113)
5.3 回归估计	(113)
5.3.1 定义	(113)
5.3.2 β 为设定常数情形	(114)
5.3.3 β 为样本回归系数情形	(115)
5.4 分层比估计与分层回归估计	(118)
5.4.1 分别比估计与联合比估计	(118)
5.4.2 分别回归估计与联合回归估计	(119)
5.4.3 各种估计量的比较与选择	(121)
* 5.5 比估计与回归估计及其方差估计的偏倚	(125)
5.5.1 有关简单随机样本中心矩阶的基本结果	(125)
5.5.2 比估计及其方差估计的偏倚	(126)
5.5.3 比估计偏倚的消除或减少办法	(128)
5.5.4 回归估计及其方差估计的偏倚	(131)
习题	(132)
第六章 不等概率抽样	(137)
6.1 概述	(137)
6.1.1 不等概率抽样的必要性及优点	(137)
6.1.2 不等概率抽样的主要分类	(138)
6.2 放回不等概率抽样	(139)

6.2.1	多项抽样与 PPS 抽样	(139)
6.2.2	多项抽样的实施方法	(139)
6.2.3	汉森—赫维茨估计量及其性质	(141)
6.2.4	数值例——职工人数调查	(143)
6.3	不放回不等概率抽样	(146)
6.3.1	包含概率与 πPS 抽样	(146)
6.3.2	霍维茨—汤普森估计量及其性质	(147)
6.3.3	$n=2$ 的严格 πPS 抽样	(148)
* 6.3.4	$n > 2$ 的严格 πPS 抽样	(151)
* 6.3.5	$n > 2$ 的两种非严格的 πPS 抽样	(152)
* 6.3.6	泊松抽样	(154)
	习题	(156)

第七章	二重抽样	(159)
7.1	概述	(159)
7.1.1	什么是二重抽样	(159)
7.1.2	二重抽样的作用	(159)
7.1.3	二步抽样估计量求均值与方差的一般公式	(160)
7.2	为分层的二重抽样	(161)
7.2.1	二重分层抽样的样本抽选方法	(161)
7.2.2	二重分层抽样的估计方法	(161)
7.3	为 PPS 抽样的二重抽样	(164)
7.4	为比估计与回归估计的二重抽样	(165)
7.4.1	二重抽样的比估计及其性质	(166)
7.4.2	二重抽样的回归估计及其性质	(168)
* 7.5	二重抽样样本量的最优分配	(170)
7.5.1	二重分层抽样样本量的最优分配	(170)
7.5.2	二重比估计与回归估计时样本量的最优分配	(171)
* 7.6	连续抽样中的样本轮换及其估计	(172)
7.6.1	样本轮换的必要性	(172)
7.6.2	样本拼配与二重回归估计的应用	(173)
	习题	(175)

第八章 整群抽样	(178)
8.1 概述	(178)
8.1.1 什么是整群抽样	(178)
8.1.2 整群抽样的特点及适用场合	(178)
8.2 群大小相等情形	(179)
8.2.1 记号	(179)
8.2.2 估计量及其性质	(180)
8.2.3 群内相关系数与设计效应	(182)
8.2.4 整群抽样效率分析及群的划分原则	(186)
8.3 估计总体比例的整群抽样	(190)
8.3.1 问题的提出	(190)
8.3.2 群大小相等情形	(190)
8.3.3 群大小不等情形	(192)
8.4 群大小不相等的一般情形	(194)
8.4.1 记号	(195)
8.4.2 按简单随机抽样抽群	(195)
8.4.3 按与群大小成比例的不等概率抽样抽群	(199)
习题	(201)

第九章 二阶及多阶抽样	(205)
9.1 概述	(205)
9.1.1 多阶抽样的定义及与其他抽样的关系	(205)
9.1.2 多阶抽样的特点和作用	(205)
9.2 初级单元大小相等时的二阶抽样	(206)
9.2.1 记号	(206)
9.2.2 总体均值 \bar{Y} 的估计量及其性质	(207)
9.2.3 关于总体比例的估计	(211)
9.2.4 最优样本量 m 与 n 的确定	(212)
9.3 初级单元大小不等时的二阶抽样(I)——对初级单元进行多项抽样情形	(214)
9.3.1 一般说明及记号	(214)
9.3.2 总体总量 Y 的估计	(215)

9.3.3 估计量是自加权的条件及对初级单元的 PPS 抽样	(216)
* 9.4 初级单元大小不等时的二阶抽样(Ⅱ)——对初级单元进行不放回抽样情形	(219)
9.4.1 用简单随机抽样抽取初级单元	(219)
9.4.2 用不放回不等概率抽样抽取初级单元	(220)
9.5 三阶及多阶抽样	(222)
9.5.1 各级单元大小相等时的三阶抽样	(222)
9.5.2 各级单元大小不等时的多阶抽样	(224)
习题	(225)
 第十章 系统抽样	(229)
10.1 概述	(229)
10.1.1 系统抽样的定义及实施方法	(229)
10.1.2 系统抽样的特点与局限性	(231)
10.2 等概率系统抽样——等距抽样	(232)
10.2.1 估计量及其性质	(232)
10.2.2 估计量方差的不同表示形式	(234)
10.2.3 估计量方差与总体单元排列顺序的关系	(238)
10.3 线性趋势总体抽样方法的改进	(240)
10.3.1 中心位置样本法与首尾校正法	(241)
10.3.2 对称系统抽样	(242)
10.4 等概率系统抽样的方差估计	(243)
10.4.1 方差估计的形式	(243)
10.4.2 各种估计量的比较及适用场合	(246)
10.5 不等概率系统抽样	(246)
10.5.1 πPS 系统抽样及其实施方法	(246)
10.5.2 估计量及其方差估计	(247)
习题	(249)
 第十一章 非抽样误差	(253)
11.1 非抽样误差的主要来源与分类	(253)
11.2 抽样框误差	(254)